

0202 030  
**RAW SEQUENCE LISTING**  
**PATENT APPLICATION US/09/003,574**

OIPE #4

DATE: 02/20/98  
TIME: 09:40:38

INPUT SET: S23591.raw

**This Raw Listing contains the General  
Information Section and up to the first 5 pages.**

## SEQUENCE LISTING

ENTERED

## (1) General Information:

(i) APPLICANT: Tripp, Cynthia Ann  
Frank, Glenn R.  
Grieve, Robert B.

(ii) TITLE OF INVENTION: NOVEL PARASITE ASTACIN  
METALLOENDOPEPTIDASE PROTEINS

(iii) NUMBER OF SEQUENCES: 36

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: SHERIDAN ROSS P.C.  
(B) STREET: 1700 LINCOLN ST., SUITE 3500  
(C) CITY: DENVER  
(D) STATE: CO  
(E) COUNTRY: USA  
(F) ZIP: 80203

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Connell, Gary J.  
(B) REGISTRATION NUMBER: 32,020  
(C) REFERENCE/DOCKET NUMBER: 2618-21-1-C1

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (303) 863-9700  
(B) TELEFAX: (303) 863-0223

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1299 base pairs

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/003,574DATE: 02/20/98  
TIME: 09:40:42

INPUT SET: S23591.raw

47 (B) TYPE: nucleic acid  
48 (C) STRANDEDNESS: single  
49 (D) TOPOLOGY: linear  
50

51 (ii) MOLECULE TYPE: cDNA  
52  
53  
54

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:  
56

57	TTTTTTTTTTT TTTTTTTTGT TTCATTGTTC AGTCAGTGGA AAATTATCGA ACGCAGAAAAG	60
58		
59	CATCACGAAA TACGTTAGAT CACATCAAAC AACTTATCAC CTTGAACGTA CAAAGAGAGA	120
60		
61	TTGGAAACAT AGATGATAAG ACATTAGCTG ATGAAATAGT ATTACAACGA CGGGATCCTG	180
62		
63	AGGCAAAATG GCATCATAAT GAACTATTCA TTAATGATCC AGATGCATAC TATCAAGGCG	240
64		
65	ATGTCGATTT GTCGGAAAAA CAAGCCGAAA TTCTAAGCGA ACATTTTAAA AATGAAATTG	300
66		
67	CTTTAACAGA GAAAGACGAC ACAATAATAC GGCGAAAAAA GAGCATTTGGT CGTGAACCAT	360
68		
69	TTTACGTAAG ATGGAATCAT AAACGTCCCA TTAGCTATGA ATTTGCGGAA AGTATTCCAT	420
70		
71	TAGAAACACG TAGAAAAATT CGTTCAGCAA TAGCAATGTG GGAAGAACGA ACATGCATAC	480
72		
73	GATTCCAAGA AAATGGCCCA AATGTAGATC GAATTGAATT TTACGACGGT GGCGGTTGTT	540
74		
75	CAAGTTTTGT CGGCCGAACA GGAGGGAATT TCAATTTCAA CACCAGGATG TGATATTATT	600
76		
77	GGTATTATAT CACATGAAAT TGGTCATACT TTAGGAATAT TTCATGAGCA AGCACGTCGT	660
78		
79	GATCAAAAAA ATCATATTTT TATTAATTAC AACAAATATC CATCAAGCCG TTGGAACAAT	720
80		
81	TTTTTTCCAT TATCAGAATA TGAAGCTGAT ATGTTTAATT TACCTTATGA TACAGGATCA	780
82		
83	GTAATGCACT ATGGTTCATA CGGATTTGCA AGAAATCCGT ATGAACCAAC TATTACAACA	840
84		
85	CGTGATAAAT TTCAACAGTA CACAATTGGG CAACGTGAAG GGCCATCATT TCTGGATTAT	900
86		
87	GCATCTGTTA AGCTTTATCT ACAAACGCAT TAATGATATT GTTATCAAAT GGATGATAAT	960
88		
89	TTCAATAAGT ATAAACAGCG CTTATCGTTG TACAGAACAA TGTGCTGATA TGCACTGCGA	1020
90		
91	TCATAATGGT TATCCGGATC CTAATAATTG CGCGAAATGC TTGTGTCCAG ATGGTTTTGC	1080
92		
93	TGGTCGTACC TGTCAATTTG TTCAATATAC ATCTTGCGGA GCTCTCATTA AGGTAAGTAT	1140
94		
95	TGTCTTTTGA CCTCTTCTCT GACTAAAAATA TAAGTTAAGC ATATGTATCT TCCGTCTAAT	1200
96		
97	GATTTTCTTG ATTTTGATTT GTTCAATGCT CTTCTTGATA ATAATATAAA AATTTTGTAA	1260
98		
99	AATAAAGTTA ACTTTTGGTC AAAAAAAAAA AAAAAAAAAA	1299

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/003,574DATE: 02/20/98  
TIME: 09:40:45

INPUT SET: S23591.raw

100  
101 (2) INFORMATION FOR SEQ ID NO:2:  
102  
103 (i) SEQUENCE CHARACTERISTICS:  
104 (A) LENGTH: 2126 base pairs  
105 (B) TYPE: nucleic acid  
106 (C) STRANDEDNESS: single  
107 (D) TOPOLOGY: linear  
108  
109 (ii) MOLECULE TYPE: cDNA  
110  
111  
112  
113 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
114  
115 GAAAGCATCA CGAAATACGT TAGATCACAT CAAACAACCTT ATCACCTTGA ACGTACAAAAG 60  
116  
117 AGAGATTGGA AACATAGATG ATAAGACATT AGCTGATGAA ATAGTATTAC AACGACGGGA 120  
118  
119 TCCTGAGGCA AAATGGCATC ATAATGAACT ATTCATTAAT GATCCAGATG CATACTATCA 180  
120  
121 AGGCGATGTC GATTTGTCGG AAAAACAAGC CGAAATTCTA AGCGAACATT TTAAAAATGA 240  
122  
123 AATTGCTTTA ACAGAGAAAAG ACGACACAAT AATACGGCGA AAAAAGAGCA TTGGTCGTGA 300  
124  
125 ACCATTTTAC GTAAGATGGA ATCATAAACG TCCATTAGC TATGAATTTG CGGAAAGTAT 360  
126  
127 TCCATTAGAA ACACGTAGAA AAATTCGTTC AGCAATAGCA ATGTGGGAAG AACGAACATG 420  
128  
129 CATACGATTC CAAGAAAATG GCCCAAATGT AGATCGAATT GAATTTTACG ACGGTGGCGG 480  
130  
131 TTGTTCAAGT TTTGTGGGCC GAACAGGAGG GAATTTCAAT TTCAACACCA GGATGTGATA 540  
132  
133 TTATTGGTAT TATATCACAT GAAATTGGTC ATACTTTAGG AATATTTTCAT GAGCAAGCAC 600  
134  
135 GTCGTGATCA AAAAAATCAT ATTTTATTA ATTACAACAA TATTCCATCA AGCCGTGGA 660  
136  
137 ACAATTTTTT TCCATTATCA GAATATGAAG CTGATATGTT TAATTTACCT TATGATACAG 720  
138  
139 GATCAGTAAT GCACTATGGT TCATACGGAT TTGCAAGAAA TCCGTATGAA CCAACTATTA 780  
140  
141 CAACACGTGA TAAATTTCAA CAGTACACAA TTGGGCAACG TGAAGGGCCA TCATTTCTGG 840  
142  
143 ATTATGCATC TGATAAACAG CGCTTATCGT TGTACAGAAC AATGTGCTGA TATGCACTGC 900  
144  
145 GATCATAATG GTTATCCGGA TCCTAATAAT TGCGCGAAAT GCTTGTGTCC AGATGGTTTT 960  
146  
147 GCTGGTCGTA CCTGTCAATT TGTTC AATAT ACATCTTGCG GAGCTCTCAT TAAGGCGAGG 1020  
148  
149 AAAATGCCTG TTACGATTTT GAGCCCAAAT TATCCAACT TCTTCAATGT TGGTGATCAA 1080  
150  
151 TGTATTTGGT TGCTTACAGC TCCACGCGTG ATTCGTAAAT TTGCAGTTTG TTGAACAATT 1140  
152

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/003,574

DATE: 02/20/98  
TIME: 09:40:49

INPUT SET: S23591.raw

```

153 TCAATTACAA TGTGAAGATA CGTGTGATAA ATCCTATGTA GAAGTGAAAG CTGACGCTGA 1200
154
155 TTTTCGACCT ACTGGATATC GATTTTGTG TTGCGGAGTG CCACGTCATA TTTTCAATC 1260
156
157 TGCGACAAAC GAGATGGTAG TAATATTTTCG CGGTTTTTGGT GATGCGGGAA ATGGCTTTAA 1320
158
159 AGCTAAAATT TGGTCAAACG TAGATGATGA TATAGCTAAT ACAATTGTAA CAACTGAAAT 1380
160
161 GGCAAAAATT TCGGAAAAAA TACCGAAGCT AACAGTTCCA ATAGTTAAAA CTATTACCAC 1440
162
163 TCCTACAATA ACAACTACTA CTGCTTTCAT GATATCACCC AAGAAAGGCA ATGTCACCGC 1500
164
165 CACGAGAGTT GCTATCACTA CTACGCCGAC TACTACAATT ACTACGACTA TTGCCGGTAC 1560
166
167 GTACCAATCA CCGTAACTAA TAATACTACA CCTGTAGTAA GTGAAACTTT ACCATCATTG 1620
168
169 CCAGTCAAGA TTCGAAACAA AATAGGTGCA TCGGAATGTG GTGAATGGAC AGAATGGACA 1680
170
171 GGTCCATGCT CTCAAGAATG TGGCGGTTGC GGAAAACGTC TTCGAACACG TCAGTGTTCA 1740
172
173 TCAGATACGG AATGTAGAAC AGAAGAAAAA CGTGCGTG TGCTTAAAGTT TGCCCATACG 1800
174
175 GGACTAATTT CCTTATCAAT AATGGAGAGT TTCATATACT TTGGAAGGGC TGCTGTGTTG 1860
176
177 GTCTATTCCG ATCGGGAGAT ATGTGTTTCA CACTTGATGA TAACGAGAAT CCATTCTCTGA 1920
178
179 AATTTCTAGA ATCACTGTTG AACATGCAAG ATTCTCGAAA AAACGATAAT TTGCCTGACT 1980
180
181 CGAAAAAGAA GTGATTGAAT GATTCGATAA TATTGATTAA TAAAACGGGT TGTATTCTCG 2040
182
183 TCATAGAGTA TCCGTTGATG TTTTATCCA AAAAAATCTC TTGCTTTTAA TTATTGTGAA 2100
184
185 TAAAACTTTT GTTTACCCAA AAAAAA 2126
186

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 191 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Cys Phe Ile Val Gln Ser Val Glu Asn Tyr Arg Thr Gln Lys Ala Ser
1           5           10           15
Arg Asn Thr Leu Asp His Ile Lys Gln Leu Ile Thr Leu Asn Val Gln
20          25          30

```

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/003,574

DATE: 02/20/98  
TIME: 09:40:52

INPUT SET: S23591.raw

```

206
207 Arg Glu Ile Gly Asn Ile Asp Asp Lys Thr Leu Ala Asp Glu Ile Val
208           35                      40                      45
209
210 Leu Gln Arg Arg Asp Pro Glu Ala Lys Trp His His Asn Glu Leu Phe
211           50                      55                      60
212
213 Ile Asn Asp Pro Asp Ala Tyr Tyr Gln Gly Asp Val Asp Leu Ser Glu
214           65                      70                      75                      80
215
216 Lys Gln Ala Glu Ile Leu Ser Glu His Phe Lys Asn Glu Ile Ala Leu
217           85                      90                      95
218
219 Thr Glu Lys Asp Asp Thr Ile Ile Arg Arg Lys Lys Ser Ile Gly Arg
220           100                     105                     110
221
222 Glu Pro Phe Tyr Val Arg Trp Asn His Lys Arg Pro Ile Ser Tyr Glu
223           115                     120                     125
224
225 Phe Ala Glu Ser Ile Pro Leu Glu Thr Arg Arg Lys Ile Arg Ser Ala
226           130                     135                     140
227
228 Ile Ala Met Trp Glu Glu Arg Thr Cys Ile Arg Phe Gln Glu Asn Gly
229           145                     150                     155                     160
230
231 Pro Asn Val Asp Arg Ile Glu Phe Tyr Asp Gly Gly Gly Cys Ser Ser
232           165                     170                     175
233
234 Phe Val Gly Arg Thr Gly Gly Asn Phe Asn Phe Asn Thr Arg Met
235           180                     185                     190
236

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

251 Ile Glu Leu Asn Phe Thr Thr Val Ala Val Val Gln Val Leu Ser Ala
252 1           5           10           15
253
254 Glu Gln Glu Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile Ile Gly
255           20           25           30
256
257 Ile Ile Ser His Glu Ile Gly His Thr Leu Gly Ile Phe His Glu Gln
258           35           40           45

```

PAGE: 1

**SEQUENCE VERIFICATION REPORT**  
**PATENT APPLICATION US/09/003,574**

DATE: 02/20/98  
TIME: 09:40:56

***INPUT SET: S23591.raw***

Line

Error

Original Text